Figure 1A
Identification of Functional Allele Families: Motivation

◆ **Repetitive Gene Systems**

- Repetitive Biochemical Functions

- E.g. systems 1, 7

◆ **Both** systems must be inactivated before problems

- E.g. (($1A_2$ or $1B_3$) and ($7A_3$ or $7C_1$))

- Automatic choice of allele combinations required

Alleles $1A_2$ or $1B_3$ May Inactivate Gene System #1	Alleles $7A_3$ or $7C_1$ May Inactivate Gene System #7

Gene System 1 (5-20 Genes)

Gene System 7 (5-20 Genes)

DNA | Allele (1 of 2) | Allele (1 of 4) | o o o | Allele (1 of 3) |

Locus A   Locus B
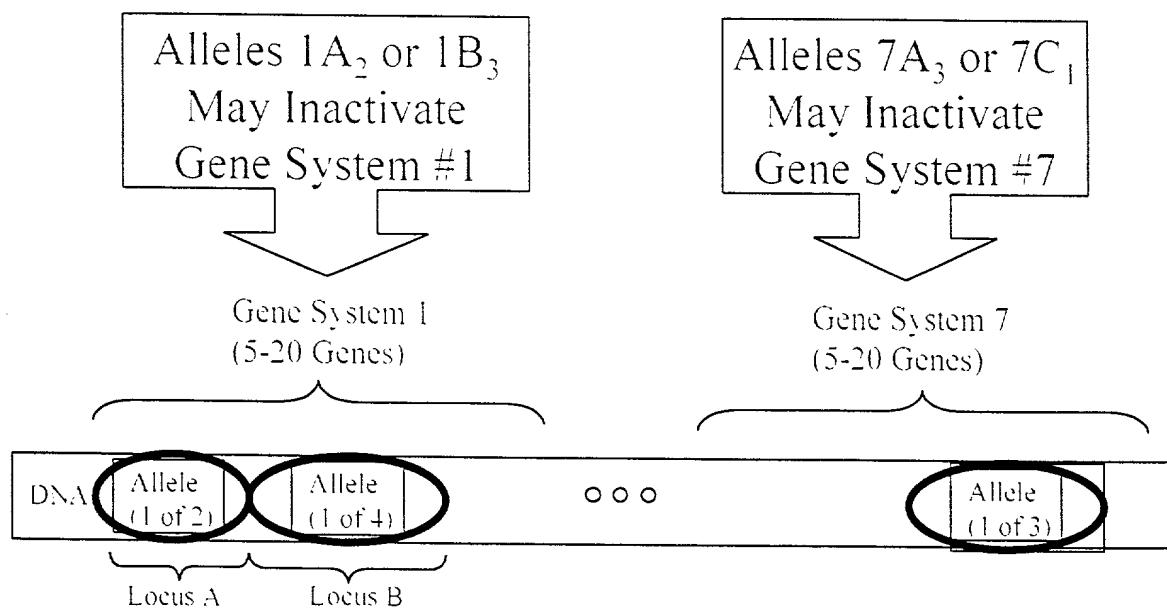
Figure 1B
Method of Identifying Clinically Relevant Allele Combinations

```
┌─────────────────────┐
│  Obtain Examples    │
│  of Inputs/Outputs  │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌─────────────────────┐
│   Run Genetic       │ ◄─────── │  Program Neural     │
│   Algorithm         │          │  Net Training Routine│
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐
│  Identify Optimal   │
│  Net Architecture   │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐
│  Report Relevant    │
│  Alleles as those   │
│  Present in Optimal  │
│  Net Architecture   │
└─────────────────────┘
```

Figure 1C
Structure of Neural Network
Training Routine

| Given Architectural Parameters (Which Inputs are Used) | Given Training Data (Sample Inputs/Outputs) |

Train Neural Net

Error Measure
for Given Architectural Parameters
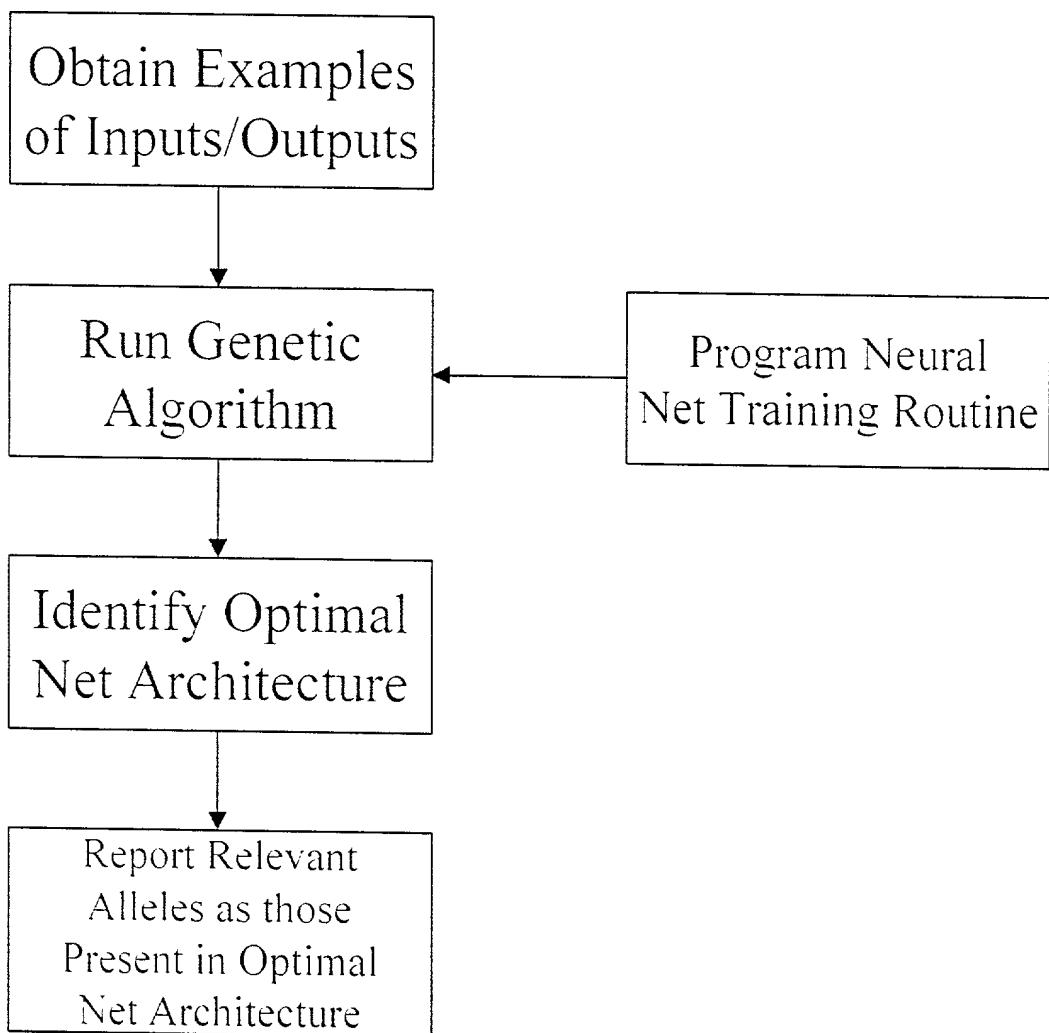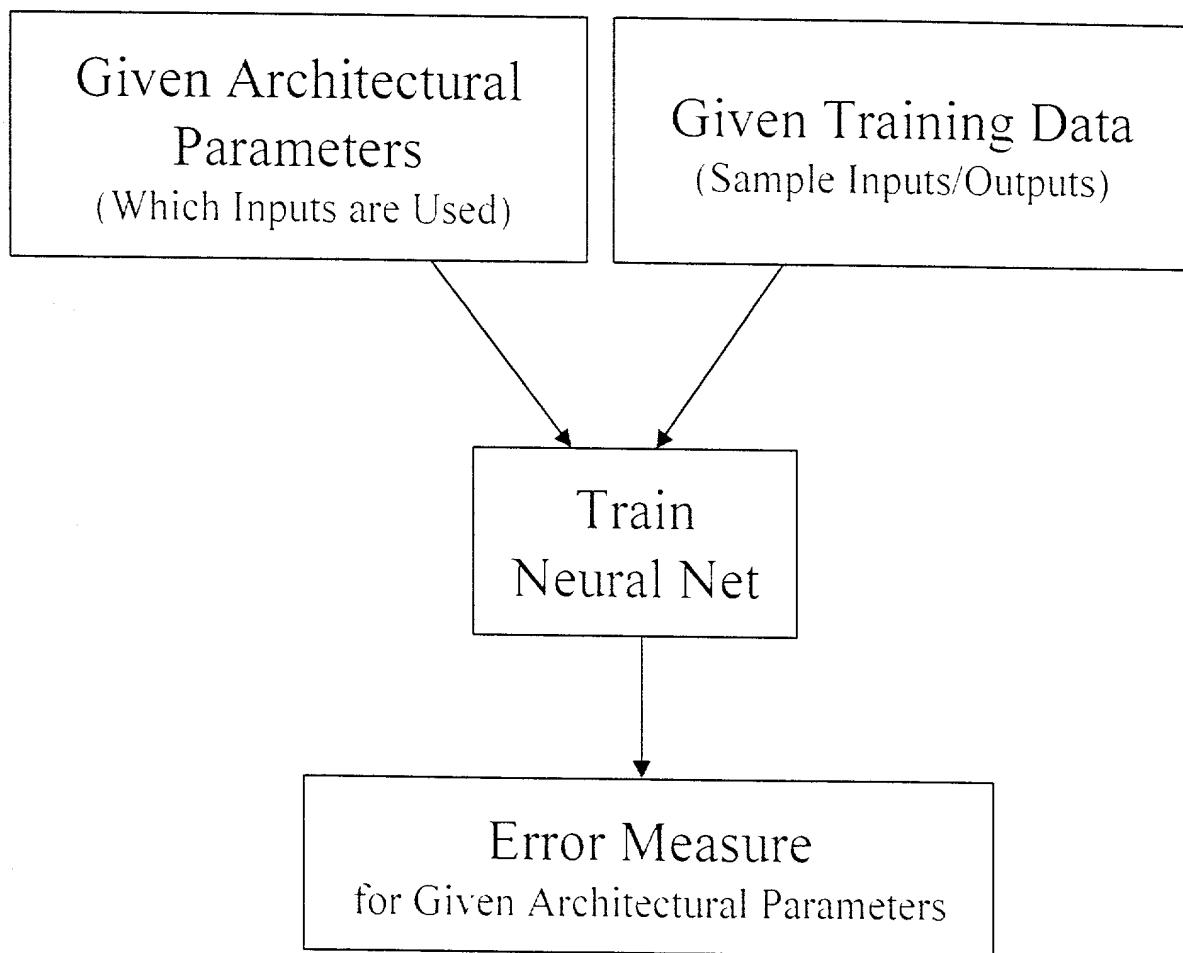
Figure 1D:
Typical Mapping Neural Network

Network Outputs

Final slab's neurons
are network outputs

Slabs of Neurons

Individual
Neurons

This convergence of inputs represents
a linear combination. in which each
input gets weighted separately.

Each neuron contains a single output
that acts as an input to the next slab

Each input synapses to
each neuron of first slab

Network Inputs
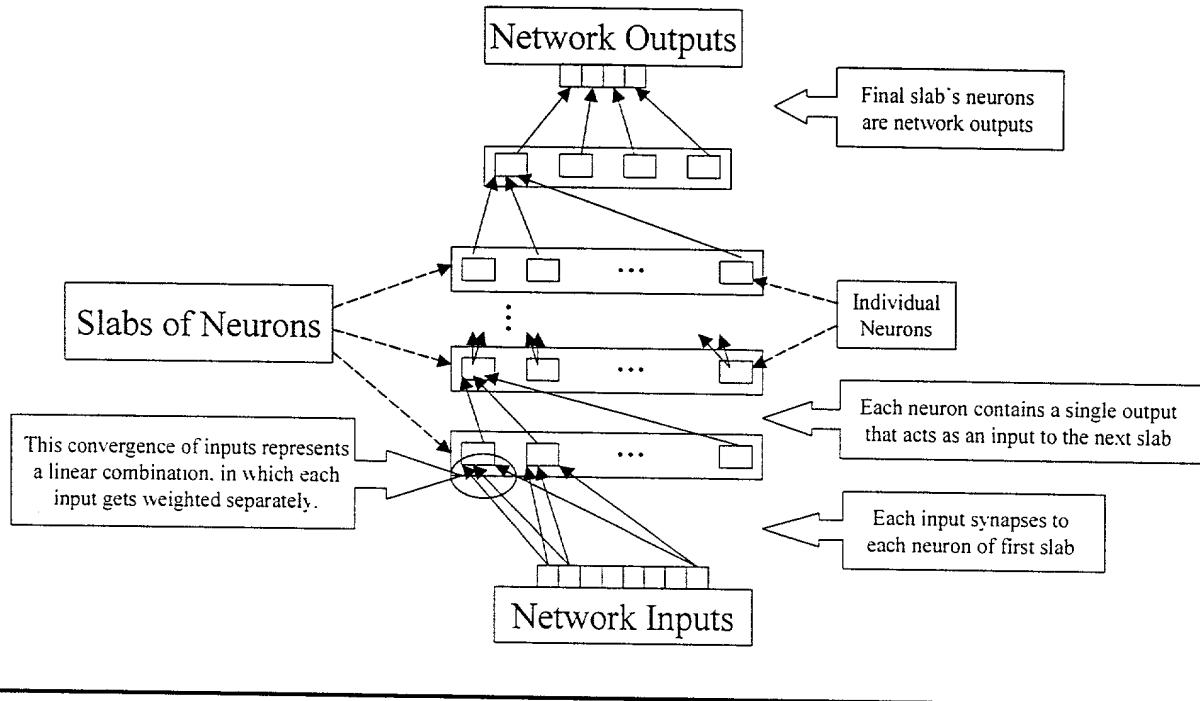
Figure 1E: Typical Genetic Algorithm

Create initial
random population

Evaluate fitness of each member

Is the stopping
criterion satisfied? ———————— Stop

Create intermediate population
using reproduction mechanism

While
(number of members
in new population) <
(population size)
do

Select two members at random

Perform crossover on members
with probability p_c

Perform mutation on members
with probability p_m

Insert members into new
population

Figure 2
Method of Predicting Clinical Variables
Given Genomic Data

Obtain Examples of Inputs/Outputs	Obtain Individual Patient's Inputs
Train Neural Net	Use Trained Net
	Report Predicted Output(s) (Range)
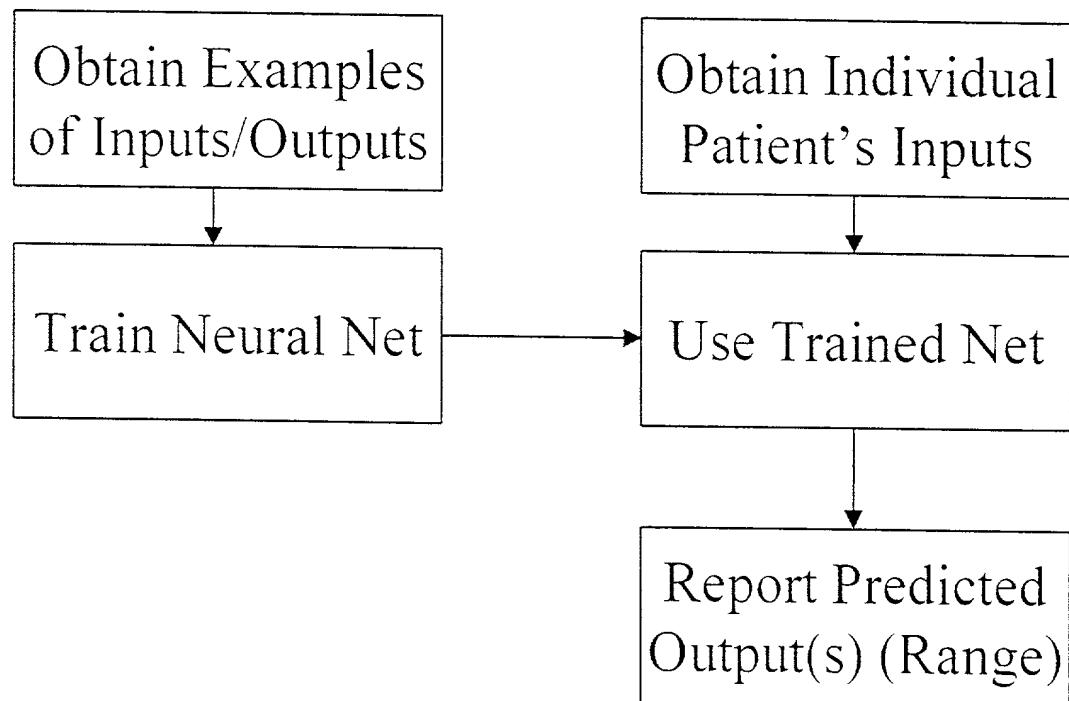
Figure 3
Genomic Methods of Screening Patients for Clinical Drug Use

Clinic

Computer

Development

Obtain Examples of Genomic Inputs

Obtain Examples of Outputs

Optimal Drug Dosage

Clinical Side Effects Measure

Train Neural Net

Use

Obtain Individual Patient's Inputs

Use Trained Net

Report Sub- (Super-) Threshold Values as Reactions

Predicted Output for Patient
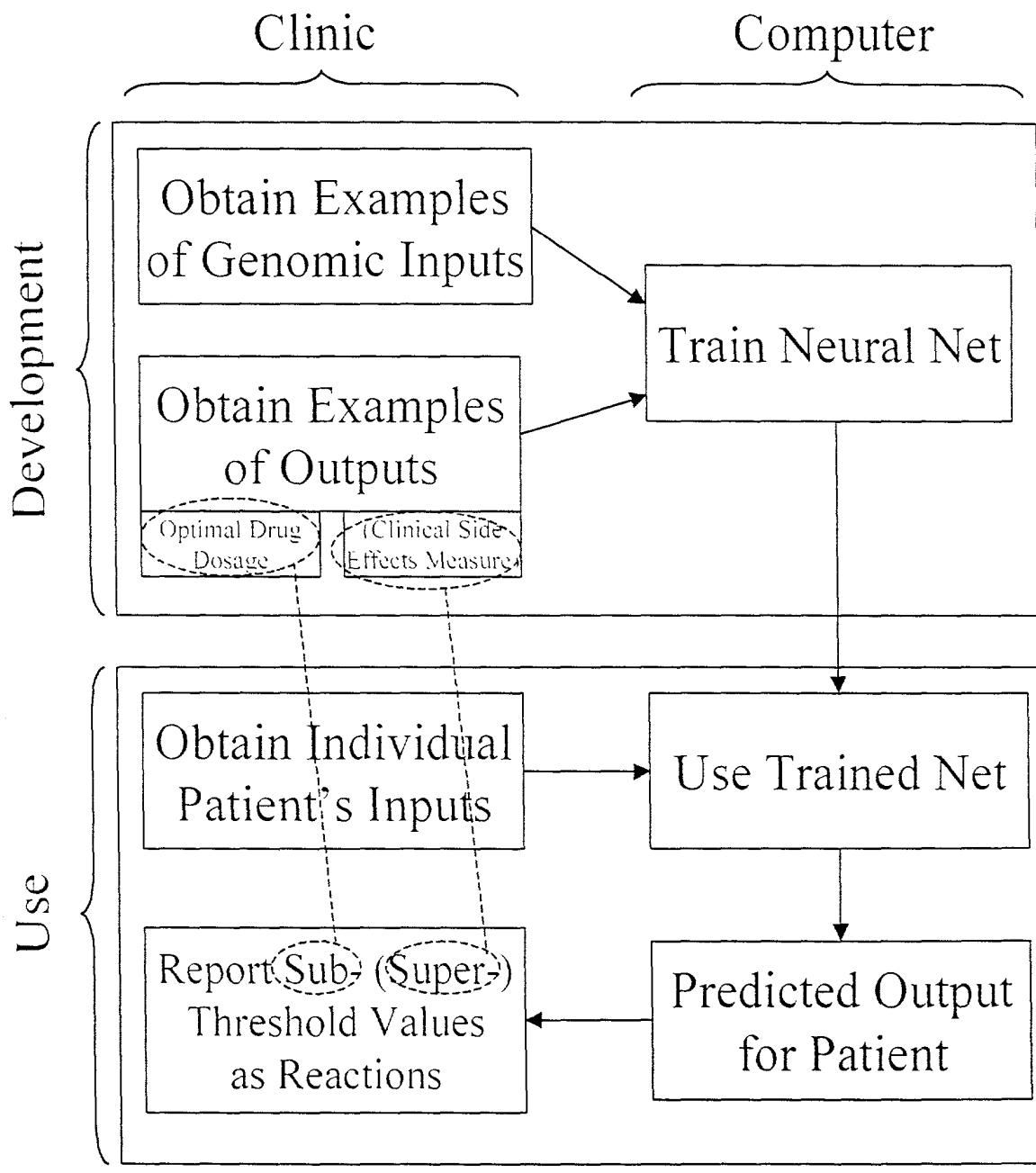
Figure 4A
GA Rolling: Illustration of Infeasible Initial Mapping Problem

Too Many Inputs	Few Outputs
Some Nearly Equivalent	

Mapping (Neural Net)

Cost Function (Error Measure)
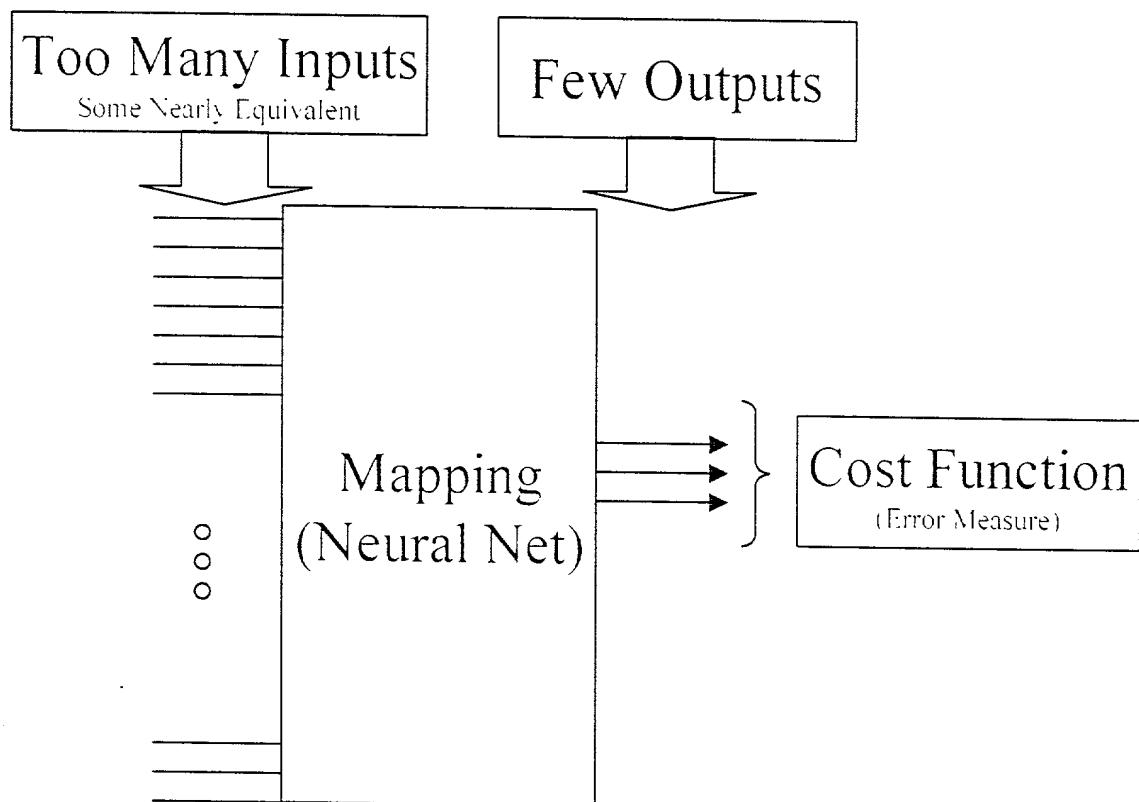
Figure 4B
GA Rolling: Illustration of Individual Category and its Genes

All Inputs	Category	Single Category
Some Nearly Equivalent: Inputs in [0.1]	Genes in [0.1]	Output in [0.1]

x 0.58
x 0.06
x 0.01
x 0.23
x 0.05
x 0.73
x 0.00

Category Preprocessor
(Apply Gene Weights)

o
o
o

x 0.07
x 0.63

Figure 4C
GA Rolling: Illustration of the Mapping
Used by the Genetic Algorithm

Fewer Input
Categories,
Each Containing
All Genomic Inputs

Same Few
Outputs

Mapping
(Neural Net)

Same Cost
Function
(Error Measure)
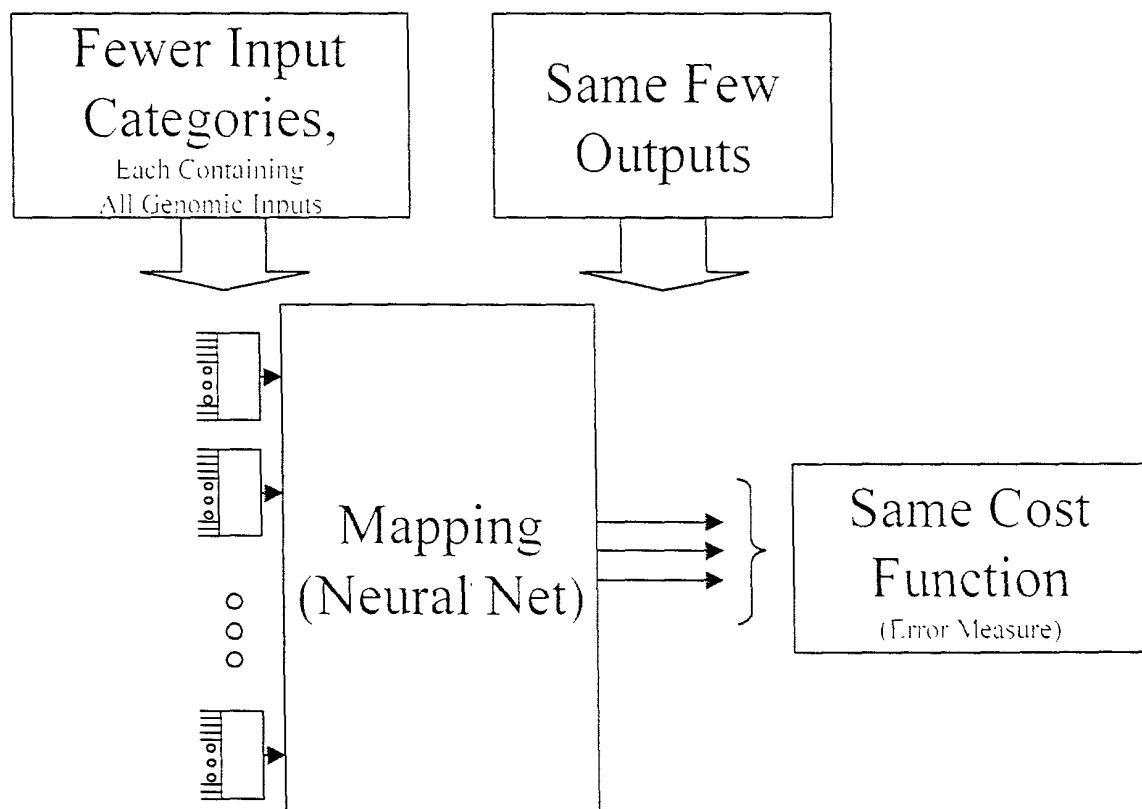
Figure 4D
GA Rolling: Illustration of the Use of the Genetic Algorithm
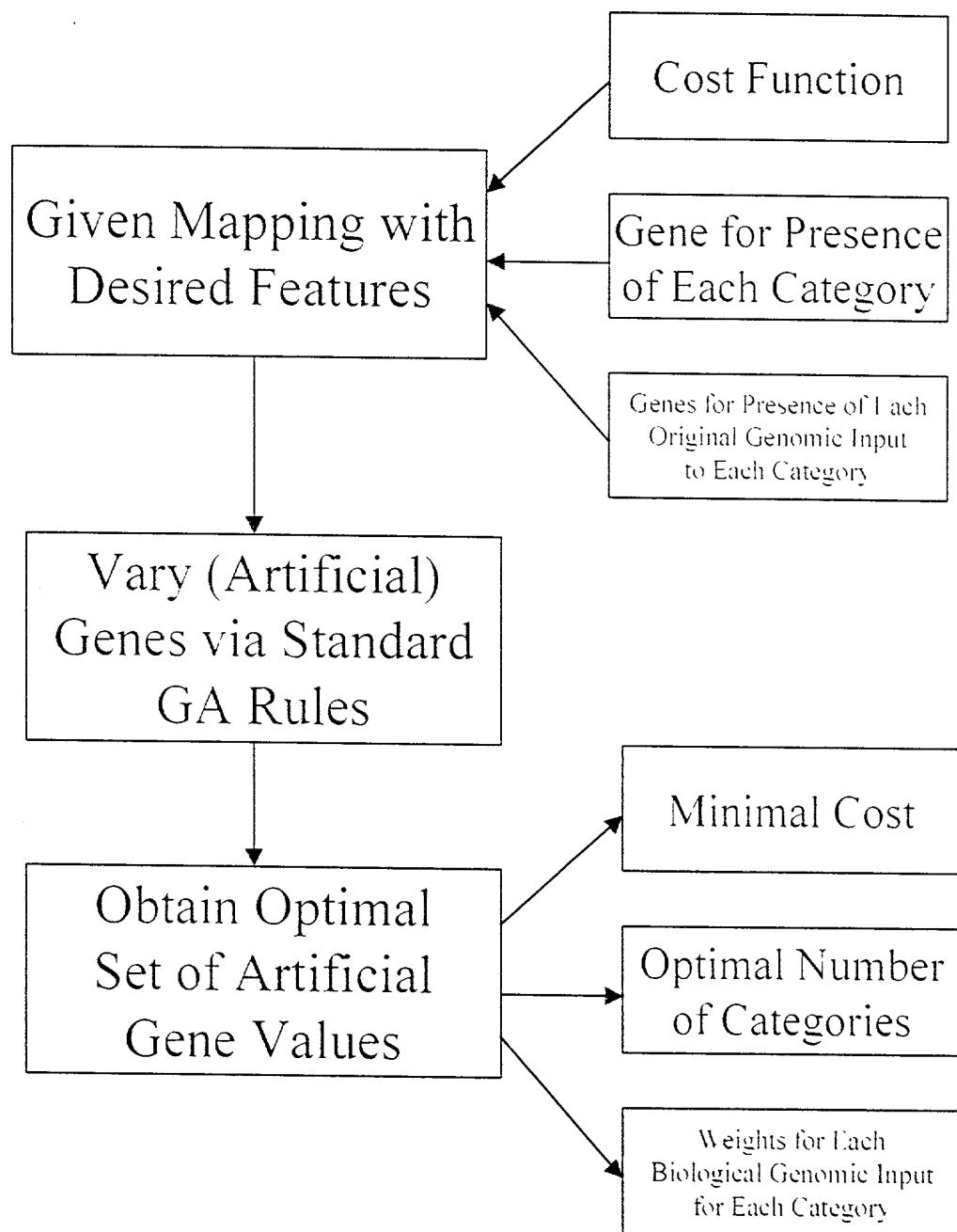
Figure 5A
Use of Functional Genomic Categorizations for Predicting Drug Interactions:
## Preliminary Constructs
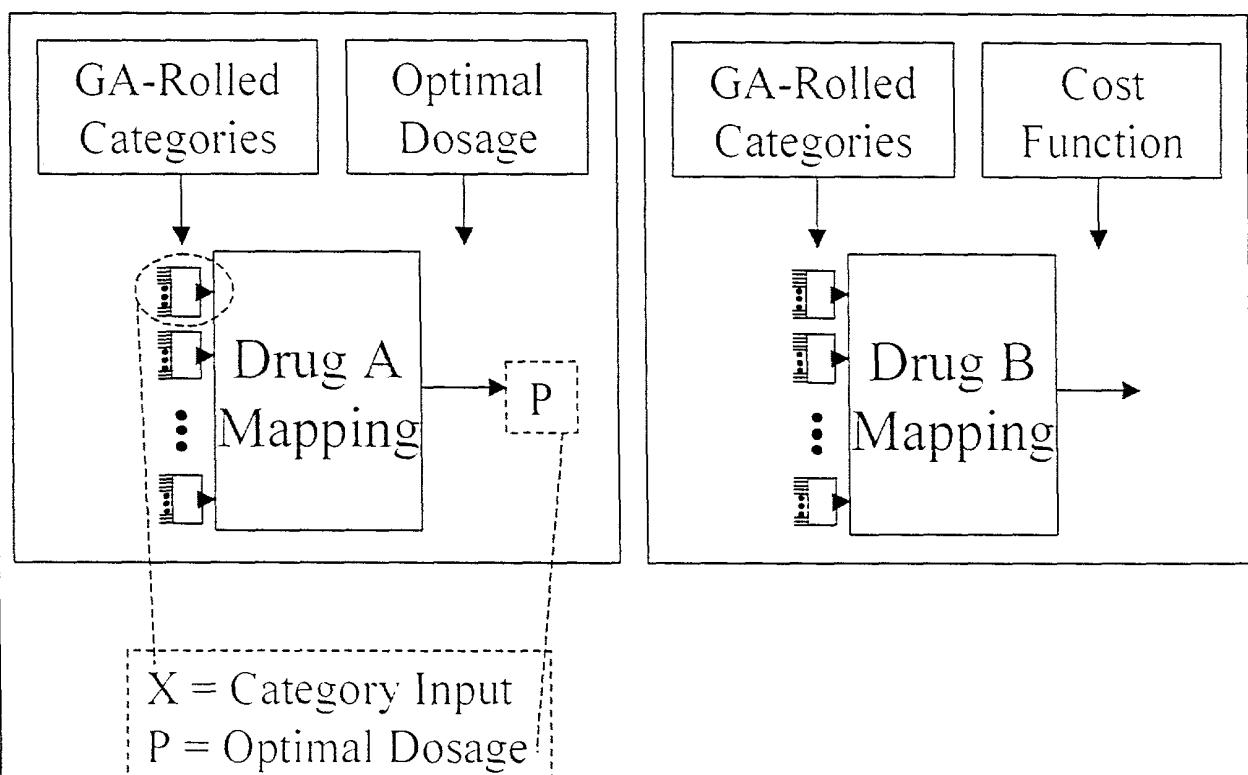
Figure 5B
Use of Functional Genomic Categorizations for Predicting Drug Interactions: Intermediate Calculations

"Equivalence" of optimal drug dosage and category X:

$$E = -\delta(\ln(P)) / \delta(\ln(X))$$

(Roughly in range [-2,2].)

Estimated equivalent of optimal drug dosage in units of category X:

$$\delta X_{drug} \sim E\ X_{patient}$$

$$(\delta X_{drug} = \text{Integral}(0, X_{patient})\ E(X')\ dX')$$

Model effect of Drug A on Drug B with effectively altered genomic inputs (to Cost B mapping):

$$X = X_{patient} + \delta X_{drug}$$

Identify corresponding increases in Cost B as interactions (of A on B) if cost increases > e.g. 30%

Figure 6A
Universal Functional Genomic Categorization:
## Assembly of Categories

Input/Output
Training Data

Given List of M
Clinical Variables
of Interest

GA Roll

Categories of
Genomic Inputs
for Variable 1

Categories of
Genomic Inputs
for Variable 2

o o o

Categories of
Genomic Inputs
for Variable M

Combine Categories
Sharing Minimum
Fraction of Super-threshold
Genomic Inputs

Universal Categories
of Genomic Inputs
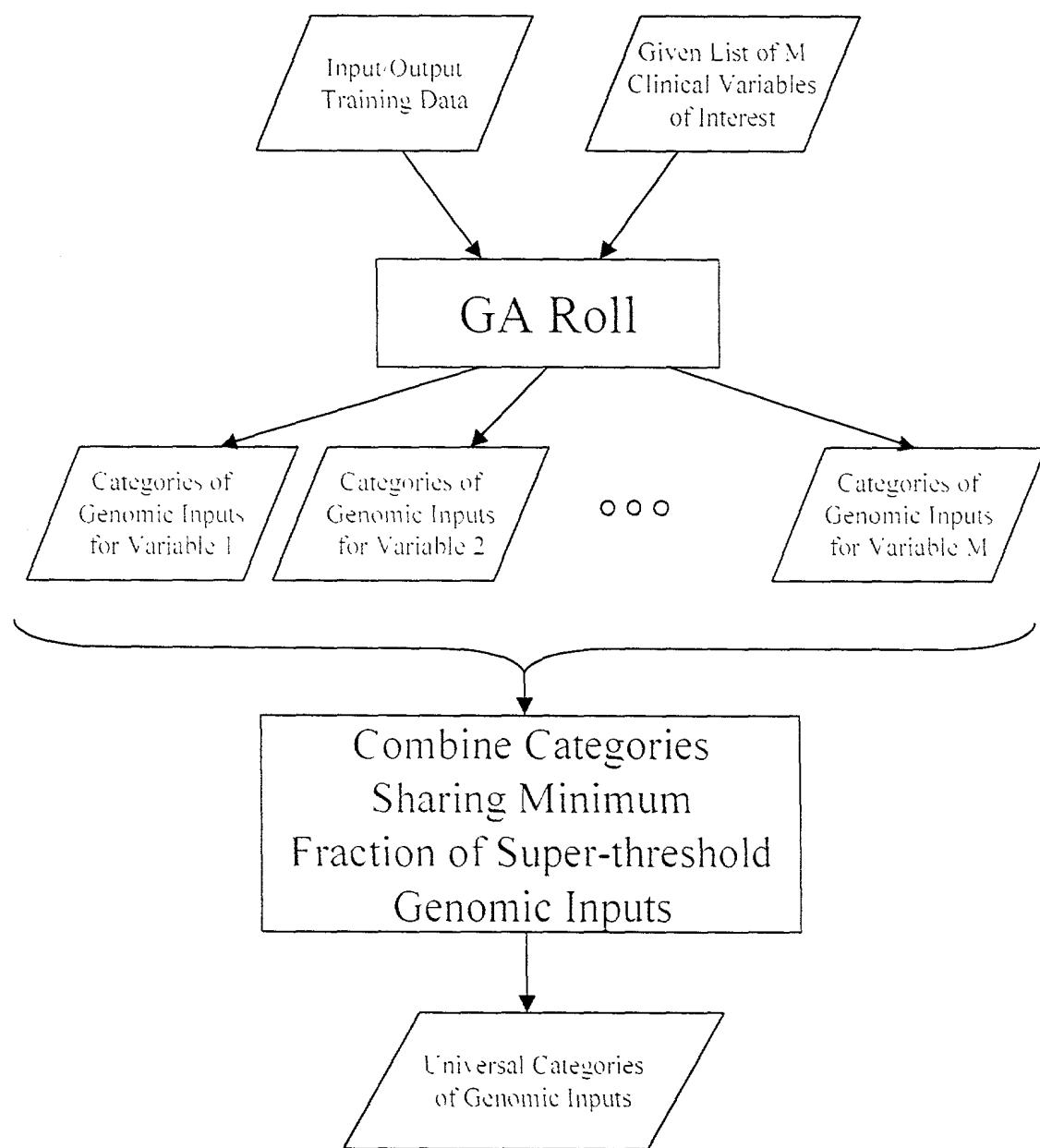
Figure 6B
Universal Functional Genomic Categorization:
# Calculation of Probabilities:
Given Information
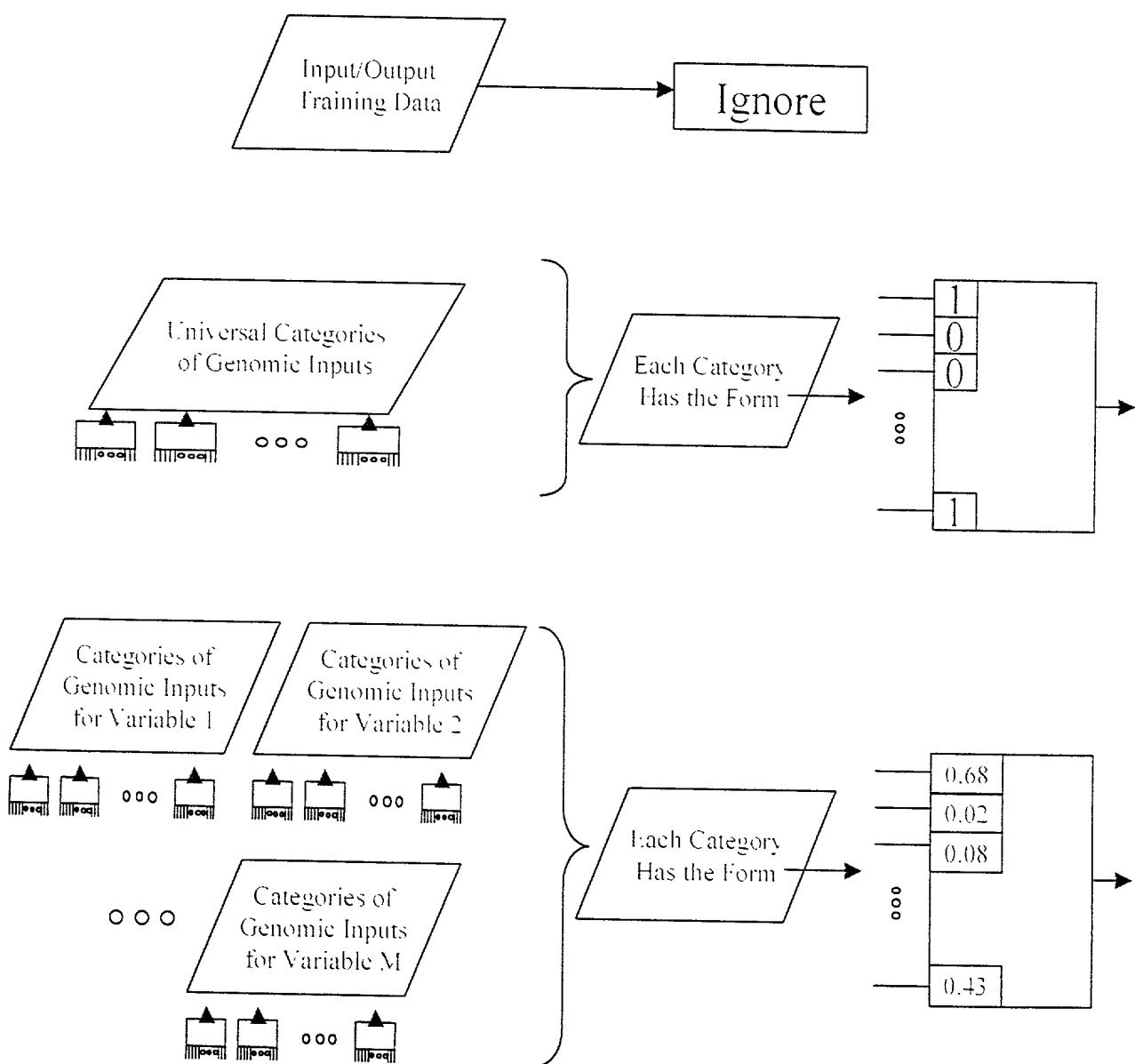
Figure 6C
Universal Functional Genomic Categorization:
Calculation of Probabilities:
Identification of Data